

LATTICE APPROXIMATIONS OF MARKOV DIFFUSIONS

YAN DOLINSKY
HEBREW UNIVERSITY OF JERUSALEM

ABSTRACT. In this note we provide a strong approximation theorem for Markov diffusion, in terms of correlated random walks. Correlated random walks lie on a grid and due to their Markov structure, dynamical programming based on such processes can be implemented efficiently. We demonstrate the applications of our strong approximation theorem to optimal stopping.

1. INTRODUCTION

One of the most popular numerical methods for stochastic control problems are tree based approximations. The standard approach to tree based approximations of diffusion processes is via the well known Euler scheme (for details see e.g. [5]). However, for a nonconstant volatility, this methodology encounters a basic difficulty. Indeed, since the volatility changes at each time, the nodes of the tree approximations do not recombine on the lattice, and this fact results in an exponential and thus a computationally explosive tree that cannot be used in many realistic situations.

The aim of this note is to present a new method which allows to approximate diffusions by discrete time and space Markov chains which lie on a grid (this is not the case for the Euler method). The grid structure allows for an efficient numerical computations of stochastic control problems via dynamical programming. Our main tool is to apply correlated random walks in order to approximate diffusion processes. A correlated random walk is a generalized random walk in the sense that the increments are not identically and independently distributed, but they only satisfy some Markov-type of conditions. The idea to use correlated random walks for approximating diffusion processes goes back to [3], where the authors studied the weak convergence of correlated random walks to Markov diffusions.

The disadvantage of the weak convergence approach is that it can not provide error estimates. In order to obtain error estimates we should consider all the processes on the same probability space, and so methods based on strong approximation theorems come into picture. In this note we extend the weak convergence results from [3] to strong approximations. The main idea is to apply the Skorokhod embedding technique for "small" perturbations of the correlated random walks. We deal with one dimensional diffusion process and assume that the diffusion coefficients are bounded and Lipschitz continuous. Moreover, we assume that the volatility is bounded away from zero. Under such assumptions we provide an error estimates

Date: October 4, 2016.

2010 Mathematics Subject Classification. 60F15, 60G40, 60J60.

Key words and phrases. Correlated random walk, Diffusion Approximations, Skorokhod embedding, Strong invariance principle, Optimal Stopping.

Research partly supported by Marie-Curie Career Integration Grant, no. 618235 .

of order $O(n^{1/4-\epsilon})$ for any $\epsilon > 0$. Finally, we demonstrate the applications of our strong approximation theorem to optimal stopping.

2. THE MAIN RESULT

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space supporting a standard Brownian motion $W = \{W_t\}_{t \geq 0}$. Let \mathcal{F}_t , $t \geq 0$ be the filtration generated by W augmented with the null sets. Consider the Markov diffusion process given by

$$(2.1) \quad dX_t = \mu(X_t)dt + \sigma(X_t)dW_t, \quad X_0 = x$$

where $x \in \mathbb{R}$ is a given initial position and $\mu, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are given functions. In this paper we assume the following assumption.

Assumption 2.1. *The functions $\mu, \sigma : \mathbb{R} \rightarrow \mathbb{R}$ are bounded and Lipschitz continuous, and $\inf_{z \in \mathbb{R}} \sigma(z) > 0$, i.e. the volatility is bounded away from zero.*

Assumption 2.1 guarantees that the SDE given by (2.1) has a unique strong solution.

The goal of this work is construct a lattice based strong approximations for $X = \{X_t\}_{t=0}^T$ where $T < \infty$ is a real number representing the time horizon. We will use similar (but not exactly the same) correlated random walks as in [3]. Our main contribution is the extension of weak approximations to strong approximations, and providing the corresponding error estimates. Before we formulate our main result, let us briefly review the main idea behind diffusion approximations (in the weak sense) by correlated random walks.

Fix $n \in \mathbb{N}$ and let $h = h(n) := \frac{T}{n}$ be the time discretization. Since the meaning is clear we will use h instead of $h(n)$. Consider the random walk

$$X_k^{(n)} = x + \sqrt{h} \sum_{i=1}^k \xi_i^{(n)}, \quad k = 0, 1, \dots, n$$

where $\xi_1^{(n)}, \dots, \xi_n^{(n)}$ are random variables with values in $\{-1, 1\}$. In the sequel, we always use the initial data $\xi_0^{(n)} \equiv 1$. Clearly, the process $\{X_k^{(n)}\}_{k=0}^n$ lies on the grid $x + \sqrt{h}\{-n, 1-n, \dots, 0, 1, \dots, n\}$.

We aim to construct a probabilistic structure so the pair $\{X_k^{(n)}, \xi_k^{(n)}\}_{k=0}^n$ forms a Markov chain weakly approximating (with the right time change) the solution of (2.1). We look for a uniformly bounded, predictable (with respect the filtration generated by $\xi_1^{(n)}, \dots, \xi_n^{(n)}$) process $\alpha^{(n)} = \{\alpha_k^{(n)}\}_{k=0}^n$ and a probability measure \mathbb{P}_n on $\sigma\{\xi_1^{(n)}, \dots, \xi_n^{(n)}\}$ such that the perturbation defined by

$$(2.2) \quad \hat{X}_k^{(n)} := X_k^{(n)} + \sqrt{h} \alpha_k^{(n)} \xi_k^{(n)}, \quad k = 0, 1, \dots, n$$

satisfies the following equations

$$(2.3) \quad \mathbb{E}_n \left(\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)} | \xi_1^{(n)}, \dots, \xi_{k-1}^{(n)} \right) = h\mu(X_{k-1}^{(n)}) + o(h), \quad k = 1, \dots, n$$

$$(2.4) \quad \mathbb{E}_n \left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2 | \xi_1^{(n)}, \dots, \xi_{k-1}^{(n)} \right) = h\sigma^2(X_{k-1}^{(n)}) + o(h), \quad k = 1, \dots, n$$

where we use the standard notation $o(h)$ to denote a random variable that converge to zero (as $h \downarrow 0$) a.s. after divided by h . We also use the convention $O(h)$ to denote a random variable that is uniformly bounded after divided by h .

From (2.2) it follows that $X_k^{(n)} - \hat{X}_k^{(n)} = o(1)$. Hence the convergence of $X^{(n)}$ is equivalent to the convergence of $\hat{X}^{(n)}$. This together with the martingale convergence Theorem 7.4.1 in [2] yields that (provided (2.3)–(2.4) hold true) $\{X_{[nt/T]}^{(n)}\}_{t=0}^T$ converge weakly to $\{X_t\}_{t=0}^T$ where $[z]$ denotes the integer part of z .

It remains to solve the equations (2.3)–(2.4). By applying (2.2)–(2.3) together with the fact that $\alpha^{(n)}$ is predictable and $(\xi^{(n)})^2 \equiv 1$ we get

$$\mathbb{E}_n \left((\hat{X}_k^{(n)} - \hat{X}_{k-1}^{(n)})^2 | \xi_1^{(n)}, \dots, \xi_{k-1}^{(n)} \right) = h(1 + 2\alpha_k^{(n)}) + h \left((\alpha_k^{(n)})^2 - (\alpha_{k-1}^{(n)})^2 \right) + o(h).$$

This together with (2.4) gives that

$$\alpha_k^{(n)} = \frac{\sigma^2(X_{k-1}^{(n)}) - 1}{2}, \quad k = 0, \dots, n$$

is a solution, where we set $X_{-1}^{(n)} \equiv x - \sqrt{h}$. Next, (2.3) yields that

$$\mathbb{E}_n \left(\xi_k^{(n)} | \xi_1^{(n)}, \dots, \xi_{k-1}^{(n)} \right) = \frac{\alpha_{k-1}^{(n)} \xi_{k-1}^{(n)} + \sqrt{h} \mu(X_{k-1}^{(n)})}{1 + \alpha_k^{(n)}}.$$

Recall that $\xi_k^{(n)} \in \{-1, 1\}$. We conclude that the probability measure \mathbb{P}_n is given by

$$\mathbb{P}_n \left(\xi_k^{(n)} = \pm 1 | \xi_1^{(n)}, \dots, \xi_{k-1}^{(n)} \right) = \frac{1}{2} \left(1 \pm \frac{\alpha_{k-1}^{(n)} \xi_{k-1}^{(n)} + \sqrt{h} \mu(X_{k-1}^{(n)})}{1 + \alpha_k^{(n)}} \right), \quad k = 1, \dots, n. \quad (2.5)$$

In view of Assumption 2.1 we assume that n is sufficiently large so \mathbb{P}_n is indeed a probability measure. Moreover, we notice that $\alpha_{k-1}^{(n)} = \frac{\sigma^2(X_{k-1}^{(n)}) - \sqrt{h} \xi_{k-1}^{(n)} - 1}{2}$. Thus the right hand side of (2.5) is determined by $X_{k-1}^{(n)}, \xi_{k-1}^{(n)}$, and so $\{X_k^{(n)}, \xi_k^{(n)}\}_{k=0}^n$ is indeed a Markov chain.

Now, we arrive to the main result of the paper.

Theorem 2.2. *It is possible to redefine $\xi_i^{(n)}$, $i \leq n$ on the Brownian probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that for any $n \in \mathbb{N}$ (sufficiently large) (2.5) holds true (where \mathbb{P}_n is replaced with \mathbb{P}) and for any $\epsilon > 0$*

$$(2.6) \quad \lim_{n \rightarrow \infty} n^{1/4-\epsilon} \mathbb{E} \left(\sup_{0 \leq t \leq T} |X_{[nt/T]}^{(n)} - X_t| \right) = 0.$$

3. PROOF OF THEOREM 2.2

Proof. Fix $n \in \mathbb{N}$. We look at the martingale part of the diffusion X . Namely, let

$$M_t := X_t - \int_0^t \mu(X_s) ds = x + \int_0^t \sigma(X_s) dW_s, \quad t \geq 0.$$

The main idea is to embed the process

$$\tilde{X}_k^{(n)} := X_k^{(n)} + \sqrt{h}(\alpha_k^{(n)} \xi_k^{(n)} - \alpha_0^{(n)} \xi_0^{(n)}) - h \sum_{i=0}^{k-1} \mu(X_i^{(n)}), \quad k = 0, 1, \dots, n$$

into the martingale $\{M_t\}_{t=0}^\infty$. Observe that $\{\tilde{X}_k^{(n)}\}_{k=0}^n$ is a martingale with respect to the measure \mathbb{P}_n , and $\tilde{X}_0^{(n)} = x$.

First Step: Redefinition of $\{X_k^{(n)}, \xi_k^{(n)}\}_{k=0}^n$ on the Brownian probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Set,

$$\alpha_0^{(n)} = \frac{\sigma^2(x - \sqrt{h}) - 1}{2}, \quad \theta_0^{(n)} = 0, \quad \xi_0^{(n)} = 1, \quad X_0^{(n)} = x.$$

For $k = 0, 1, \dots, n-1$ define by recursion the following random variables

$$(3.1) \quad \alpha_{k+1}^{(n)} = \frac{\sigma^2(X_k^{(n)}) - 1}{2},$$

$$\theta_{k+1}^{(n)} = \inf \left\{ t > \theta_k^{(n)} : |M_t - M_{\theta_k^{(n)}} + \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} + h\mu(X_k^{(n)})| = \sqrt{h}(1 + \alpha_{k+1}^{(n)}) \right\},$$

$$(3.2) \quad \xi_{k+1}^{(n)} = \mathbb{I}_{\theta_{k+1}^{(n)} < \infty} \operatorname{sgn} \left(M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}} + \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} + h\mu(X_k^{(n)}) \right),$$

where we put $\operatorname{sgn}(z) = 1$ for $z > 0$ and -1 otherwise. Finally, set

$$(3.3) \quad X_{k+1}^{(n)} = X_k^{(n)} + \sqrt{h}\xi_{k+1}^{(n)}.$$

Observe that $\theta_{k+1}^{(n)}$ is the stopping time which corresponds to the Skorokhod embedding of the binary random variable with values in the (random) set $\{\pm\sqrt{h}(1 + \alpha_{k+1}^{(n)}) - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)})\}$, into the martingale $\{M_t - M_{\theta_k^{(n)}}\}_{t \geq \theta_k^{(n)}}$.

Second Step: In this step we derive the basic properties of the construction. We show by induction that for any $k = 0, 1, \dots, n$ we have $\theta_k^{(n)} < \infty$ and (2.5) holds true (where \mathbb{P}_n replaced by \mathbb{P}). For $k = 0$ the statement is trivial. Assume it is correct for $1, 2, \dots, k$, let us prove it for $k+1$. From the induction assumption we have $|X_k^{(n)} - X_{k-1}^{(n)}| = \sqrt{h}$. This together with the fact that σ is a bounded and Lipschitz continuous gives $|\alpha_{k+1}^{(n)} - \alpha_k^{(n)}| = O(\sqrt{h})$. Since σ is bounded away from zero then $\alpha^{(n)}$ is bounded away (uniformly) from $-1/2$. Hence, (recall that μ is bounded) for sufficiently large n we have

$$-\sqrt{h}(1 + \alpha_{k+1}^{(n)}) - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)}) < 0 < \sqrt{h}(1 + \alpha_{k+1}^{(n)}) - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)}).$$

Thus by using the fact that volatility of the martingale M is bounded away from zero we conclude that $\theta_{k+1}^{(n)} - \theta_k^{(n)} < \infty$ as required.

Next, we establish (2.5). From (3.1) we get

$$(3.4) \quad M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}} = \sqrt{h}(1 + \alpha_{k+1}^{(n)})\xi_{k+1}^{(n)} - \sqrt{h}\alpha_k^{(n)}\xi_k^{(n)} - h\mu(X_k^{(n)}).$$

The stochastic process $\{M_t - M_{\theta_k^{(n)}}\}_{t=\theta_k^{(n)}}^{\theta_{k+1}^{(n)}}$ is a bounded martingale and so $\mathbb{E}(M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}} | \mathcal{F}_{\theta_k^{(n)}}) = 0$. Hence, from (3.4)

$$\mathbb{E}(\xi_{k+1}^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) = \frac{\alpha_k^{(n)}\xi_k^{(n)} + \sqrt{h}\mu(X_k^{(n)})}{1 + \alpha_{k+1}^{(n)}}.$$

Since $\xi_{k+1}^{(n)} \in \{-1, 1\}$ we get

$$(3.5) \quad \mathbb{P}(\xi_{k+1}^{(n)} = \pm 1 | \mathcal{F}_{\theta_k^{(n)}}) = \frac{1}{2} \left(1 \pm \frac{\alpha_k^{(n)}\xi_k^{(n)} + \sqrt{h}\mu(X_k^{(n)})}{1 + \alpha_{k+1}^{(n)}} \right)$$

and conclude that (the above right hand side is $\sigma\{\xi_1^{(n)}, \dots, \xi_k^{(n)}\}$ measurable)

$$(3.6) \quad \mathbb{P}(\xi_{k+1}^{(n)} = \pm 1 | \xi_1^{(n)}, \dots, \xi_k^{(n)}) = \frac{1}{2} \left(1 \pm \frac{\alpha_k^{(n)} \xi_k^{(n)} + \sqrt{h} \mu(X_k^{(n)})}{1 + \alpha_{k+1}^{(n)}} \right)$$

as well. This completes the second step.

Third Step: In this step we derive some elementary estimates and apply the discrete version of the Gronwall inequality. Recall the definition of $\hat{X}^{(n)}$ given by (2.2). From (3.4) it follows that

$$(3.7) \quad M_{\theta_k^{(n)}} = \hat{X}_k^{(n)} - \sqrt{h} \alpha_0^{(n)} - h \sum_{i=0}^{k-1} \mu(X_i^{(n)}), \quad k = 0, 1, \dots, n.$$

Since the process $\alpha^{(n)}$ is uniformly bounded and μ is Lipschitz continuous, then from (2.2) and (3.7) we obtain

$$(3.8) \quad \begin{aligned} |X_k^{(n)} - X_{\theta_k^{(n)}}| &= \\ &O(\sqrt{h}) + \left| h \sum_{i=0}^{k-1} \mu(X_i^{(n)}) - \sum_{i=0}^{k-1} \int_{\theta_{i+1}^{(n)}}^{\theta_i^{(n)}} \mu(X_t) dt \right| = \\ &O(\sqrt{h}) + \left| h \sum_{i=0}^{k-1} \mu(X_i^{(n)}) - \sum_{i=0}^{k-1} \mu(X_{\theta_i^{(n)}}) (\theta_{i+1}^{(n)} - \theta_i^{(n)}) \right| \\ &\quad + O(1) \sum_{i=0}^{k-1} \sup_{\theta_i^{(n)} \leq t \leq \theta_{i+1}^{(n)}} |X_t - X_{\theta_i^{(n)}}| (\theta_{i+1}^{(n)} - \theta_i^{(n)}) = \\ &O(\sqrt{h}) + O(1) \sum_{i=0}^{k-1} L_i + O(h) \sum_{i=0}^{k-1} |X_i^{(n)} - X_{\theta_i^{(n)}}| + \left| \sum_{i=0}^{k-1} (I_i + J_i) \right| \end{aligned}$$

where

$$I_i := \mu(X_{\theta_i^{(n)}}) \left(\theta_{i+1}^{(n)} - \theta_i^{(n)} - \mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)} | \mathcal{F}_{\theta_i^{(n)}}) \right)$$

$$J_i := \mu(X_{\theta_i^{(n)}}) \left(\mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)} | \mathcal{F}_{\theta_i^{(n)}}) - h \right)$$

$$L_i = \sup_{\theta_i^{(n)} \leq t \leq \theta_{i+1}^{(n)}} |X_t - X_{\theta_i^{(n)}}| (\theta_{i+1}^{(n)} - \theta_i^{(n)}).$$

Next, from (3.1) and the Ito isometry it follows

$$(3.9) \quad \begin{aligned} \mathbb{E} \left(\int_{\theta_k^{(n)}}^{\theta_{k+1}^{(n)}} \sigma^2(X_t) dt | \mathcal{F}_{\theta_k^{(n)}} \right) &= \mathbb{E} \left((M_{\theta_{k+1}^{(n)}} - M_{\theta_k^{(n)}})^2 | \mathcal{F}_{\theta_k^{(n)}} \right) = \\ &h(1 + 2\alpha_{k+1}^{(n)}) + h \left((\alpha_{k+1}^{(n)})^2 - (\alpha_k^{(n)})^2 \right) + O(h^{3/2}) = h\sigma^2(X_k^{(n)}) + O(h^{3/2}). \end{aligned}$$

On the other hand, the function σ is bounded and Lipschitz, and so

$$(3.10) \quad \begin{aligned} \mathbb{E} \left(\int_{\theta_k^{(n)}}^{\theta_{k+1}^{(n)}} \sigma^2(X_t) dt | \mathcal{F}_{\theta_k^{(n)}} \right) &= \sigma^2(X_k^{(n)}) \mathbb{E}(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) + \\ &O(1) \mathbb{E} \left(\sup_{\theta_k^{(n)} \leq t \leq \theta_{k+1}^{(n)}} |X_t - X_{\theta_k^{(n)}}| (\theta_{k+1}^{(n)} - \theta_k^{(n)}) | \mathcal{F}_{\theta_k^{(n)}} \right) + \\ &O(1) \left| X_k^{(n)} - X_{\theta_k^{(n)}} \right| \mathbb{E} \left(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}} \right). \end{aligned}$$

From (3.9)–(3.10) and the fact that σ bounded away from zero we get

$$(3.11) \quad \begin{aligned} \mathbb{E}(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) &= h + O(h^{3/2}) + \\ O(1) \mathbb{E} \left(\sup_{\theta_k^{(n)} \leq t \leq \theta_{k+1}^{(n)}} |X_t - X_{\theta_k^{(n)}}| (\theta_{k+1}^{(n)} - \theta_k^{(n)}) | \mathcal{F}_{\theta_k^{(n)}} \right) + \\ O(1) \left| X_k^{(n)} - X_{\theta_k^{(n)}} \right| \mathbb{E} \left(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}} \right). \end{aligned}$$

Clearly, (3.9) implies that (σ is bounded away from zero) $\mathbb{E}(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) = O(h)$. This together with (3.11) gives (μ is bounded)

$$(3.12) \quad |J_i| = O(h^{3/2}) + O(h) |X_i^{(n)} - X_{\theta_i^{(n)}}| + O(1) \mathbb{E}(L_i | \mathcal{F}_{\theta_i^{(n)}}).$$

From (3.8) and (3.12) we conclude that for any $k = 1, \dots, n$

$$\begin{aligned} \max_{0 \leq j \leq k} |X_j^{(n)} - X_{\theta_j^{(n)}}| &= \\ O(\sqrt{h}) + O(h) \sum_{i=0}^{k-1} \max_{0 \leq j \leq i} |X_j^{(n)} - X_{\theta_j^{(n)}}| + \\ \max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| + O(1) \sum_{i=0}^{n-1} L_i + O(1) \sum_{i=0}^{n-1} \mathbb{E}(L_i | \mathcal{F}_{\theta_i^{(n)}}). \end{aligned}$$

Thus,

$$\begin{aligned} \mathbb{E} \left(\max_{0 \leq j \leq k} |X_j^{(n)} - X_{\theta_j^{(n)}}| \right) &= \\ O(\sqrt{h}) + O(h) \sum_{i=0}^{k-1} \mathbb{E} \left(\max_{0 \leq j \leq i} |X_j^{(n)} - X_{\theta_j^{(n)}}| \right) \\ + \mathbb{E} \left(\max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| \right) + O(1) \sum_{i=0}^{n-1} \mathbb{E}(L_i). \end{aligned}$$

Finally, from the discrete version of Gronwall's inequality (see [1])

$$(3.13) \quad \begin{aligned} \mathbb{E} \left(\max_{0 \leq i \leq n} |X_i^{(n)} - X_{\theta_i^{(n)}}| \right) &= (1 + O(h))^n \times \\ \left(O(\sqrt{h}) + \mathbb{E} \left(\max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| \right) + O(1) \sum_{i=0}^{n-1} \mathbb{E}(L_i) \right) \\ &= O(\sqrt{h}) + O(1) \mathbb{E} \left(\max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| \right) + O(1) \sum_{i=0}^{n-1} \mathbb{E}(L_i). \end{aligned}$$

Fourth Step: In this step we estimate $\mathbb{E} \left(\max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| \right)$ and $\mathbb{E}(L_i | \mathcal{F}_{\theta_i^{(n)}})$, $i = 0, 1, \dots, n-1$. By applying the Burkholder-Davis-Gundy inequality for the martingale $\{M_t - M_{\theta_k^{(n)}}\}_{t=\theta_k^{(n)}}^{\theta_{k+1}^{(n)}}$ it follows that

$$\begin{aligned} \mathbb{E} \left(\left(\int_{\theta_k^{(n)}}^{\theta_{k+1}^{(n)}} \sigma^2(X_t) dt \right)^m \middle| \mathcal{F}_{\theta_k^{(n)}} \right) &= \\ O(1) \mathbb{E} \left(\max_{\theta_k^{(n)} \leq t \leq \theta_{k+1}^{(n)}} \left(M_t - M_{\theta_k^{(n)}} \right)^{2m} \middle| \mathcal{F}_{\theta_k^{(n)}} \right) &= O(h^m), \quad m > 1/2 \end{aligned}$$

where the last equality follows from the fact that $\max_{\theta_k^{(n)} \leq t \leq \theta_{k+1}^{(n)}} |M_t - M_{\theta_k^{(n)}}| = O(\sqrt{h})$. Since σ is bounded away from zero we get

$$(3.14) \quad \mathbb{E} \left((\theta_{k+1}^{(n)} - \theta_k^{(n)})^m | \mathcal{F}_{\theta_k^{(n)}} \right) = O(h^m), \quad m > 1/2.$$

Next, observe that $\sum_{i=0}^k I_i$, $k = 0, \dots, n-1$ is a martingale. From the Doob–Kolmogorov inequality, (3.14) and the fact that μ is bounded we conclude

$$\mathbb{E} \left(\max_{0 \leq k \leq n-1} \left(\sum_{i=0}^k I_i \right)^2 \right) \leq 4\mathbb{E} \left(\sum_{i=0}^{n-1} I_i^2 \right) = 4\mathbb{E} \left(\sum_{i=0}^{n-1} \mu^2(X_{\theta_i^{(n)}}) O(h^2) \right) = O(h).$$

Thus

$$(3.15) \quad \mathbb{E} \left(\max_{0 \leq k \leq n-1} \left| \sum_{i=0}^k I_i \right| \right) = O(\sqrt{h}).$$

Finally, we estimate $\mathbb{E}(L_i | \mathcal{F}_{\theta_i^{(n)}})$ for $i = 0, 1, \dots, n-1$. Clearly,

$$\begin{aligned} & \sup_{\theta_i^{(n)} \leq t \leq \theta_{i+1}^{(n)}} |X_t - X_{\theta_i^{(n)}}| \leq \\ & \sup_{\theta_i^{(n)} \leq t \leq \theta_{i+1}^{(n)}} |M_t - M_{\theta_i^{(n)}}| + \|\mu\|_\infty (\theta_{i+1}^{(n)} - \theta_i^{(n)}) \\ & = O(\sqrt{h}) + \|\mu\|_\infty (\theta_{i+1}^{(n)} - \theta_i^{(n)}). \end{aligned}$$

Hence, from (3.14) we get

$$(3.16) \quad \mathbb{E}(L_i | \mathcal{F}_{\theta_i^{(n)}}) \leq \mathbb{E} \left(O(\sqrt{h}) (\theta_{i+1}^{(n)} - \theta_i^{(n)}) + \|\mu\|_\infty (\theta_{i+1}^{(n)} - \theta_i^{(n)})^2 | \mathcal{F}_{\theta_i^{(n)}} \right) = O(h^{3/2}).$$

This together with (3.13) and (3.15) yields

$$(3.17) \quad \mathbb{E} \left(\max_{0 \leq i \leq n} |X_i^{(n)} - X_{\theta_i^{(n)}}| \right) = O(\sqrt{h}).$$

Fifth Step: In this step we complete the proof. From the Burkholder–Davis–Gundy inequality and the trivial inequality $(a+b)^m \leq 2^m(a^m + b^m)$, $a, b \geq 0$, $m > 0$ it follows that for any $m > 1$ and stopping times $\Theta_1 \leq \Theta_2$

$$\begin{aligned} (3.18) \quad & \mathbb{E} \left(\sup_{\Theta_1 \leq t \leq \Theta_2} |X_t - X_s|^m \right) \leq \\ & 2^m \mathbb{E} \left(\sup_{\Theta_1 \leq t \leq \Theta_2} |M_t - M_{\Theta_1}|^m + \|\mu\|_\infty^m (\Theta_2 - \Theta_1)^m \right) = \\ & O(1) \mathbb{E} \left(\left| \int_{\Theta_1}^{\Theta_2} \sigma^2(X_t) dt \right|^{m/2} + (\Theta_2 - \Theta_1)^m \right) = \\ & O(1) \mathbb{E} \left((\Theta_2 - \Theta_1)^{m/2} + (\Theta_2 - \Theta_1)^m \right). \end{aligned}$$

Thus, by applying this statement for $\Theta_1 = \theta_k^{(n)} \wedge kh$, $\Theta_2 = \theta_k^{(n)} \vee kh$ and $\Theta'_1 = kh$, $\Theta'_2 = (k+1)h$ for $k = 0, 1, \dots, n-1$ we obtain

$$\begin{aligned} (3.19) \quad & \mathbb{E} \left(\max_{kh \leq t \leq (k+1)h} |X_t - X_{\theta_k^{(n)}}|^m \right) \leq \\ & 2^m \mathbb{E} \left(\max_{\Theta_1 \leq t \leq \Theta_2} |X_t - X_{\Theta_1}|^m \right) + 2^m \mathbb{E} \left(\max_{\Theta'_1 \leq t \leq \Theta'_2} |X_t - X_{\Theta'_1}|^m \right) \\ & = O(1) \mathbb{E} \left(|\theta_k^{(n)} - kh|^{m/2} + |\theta_k^{(n)} - kh|^m \right) + O(1) h^{m/2}, \quad m > 1. \end{aligned}$$

Next, introduce the martingale

$$N_k = \sum_{i=0}^k \left(\theta_{i+1}^{(n)} - \theta_i^{(n)} - \mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)} | \mathcal{F}_{\theta_i^{(n)}}) \right), \quad k = 0, 1, \dots, n-1.$$

From (3.11), (3.14) and (3.16)–(3.17) it follows that $\mathbb{E}(\theta_{k+1}^{(n)} - \theta_k^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) = h + O(h^{3/2})$. Hence, for any $m > 1$

$$\mathbb{E}(|\theta_k^{(n)} - kh|^m) \leq \mathbb{E}((|N_k| + O(\sqrt{h}))^m) = O(1)h^{m/2} + O(1)\mathbb{E}(|N_k|^m) =$$

by applying the Burkholder–Davis–Gundy inequality

$$\begin{aligned} &= O(1)h^{m/2} + O(1)\mathbb{E}\left(\left(\sum_{i=0}^k (\theta_{i+1}^{(n)} - \theta_i^{(n)} - \mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)} | \mathcal{F}_{\theta_i^{(n)}}))^2\right)^{m/2}\right) \\ &= O(1)h^{m/2} + O(1)n^{m/2} \sum_{i=0}^k \mathbb{E}\left(|\theta_{i+1}^{(n)} - \theta_i^{(n)} - \mathbb{E}(\theta_{i+1}^{(n)} - \theta_i^{(n)} | \mathcal{F}_{\theta_i^{(n)}})|^m\right) = \\ &\quad O(1)h^{m/2} + O(1)n^{m/2+1}2^m O(h^m) \end{aligned}$$

where the last equality follows from (3.14) and the inequality $|a-b|^m \leq 2^m(a^m+b^m)$, $a, b \geq 0$. Thus

$$(3.20) \quad \mathbb{E}(|\theta_k^{(n)} - kh|^m) = O(h^{m/2-1}), \quad m > 1, \quad k = 0, 1, \dots, n.$$

From (3.19)–(3.20) we conclude that for any $m > 2$

$$\begin{aligned} &\mathbb{E}\left(\max_{0 \leq k \leq n-1} \max_{kh \leq t \leq (k+1)h} |X_t - X_{\theta_k^{(n)}}|^m\right) \leq \\ &\sum_{0 \leq k \leq n-1} \mathbb{E}\left(\max_{kh \leq t \leq (k+1)h} |X_t - X_{\theta_k^{(n)}}|^m\right) = \\ &\quad nO(h^{m/4} - 1) = O(h^{m/4-2}). \end{aligned}$$

Finally, from the Jensen inequality and (3.17)

$$\begin{aligned} &\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_{[nt/T]}^{(n)} - X_t|\right) \leq \mathbb{E}\left(\max_{0 \leq i \leq n} |X_i^{(n)} - X_{\theta_i^{(n)}}|\right) + \\ &\mathbb{E}\left(\max_{0 \leq k \leq n-1} \max_{kh \leq t \leq (k+1)h} |X_t - X_{\theta_k^{(n)}}|\right) \leq \\ &\quad = O(\sqrt{h}) + (O(h^{m/4-2}))^{1/m} = O(h^{1/4-2/m}). \end{aligned}$$

Since $m > 2$ was arbitrary we complete the proof of (2.6). \square

Remark 3.1. *An interesting open question is to find a sharp estimate for*

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_{[nt/T]}^{(n)} - X_t|\right).$$

For the case where X is the standard Brownian motion (i.e. $\sigma \equiv 1$ and $\mu \equiv 0$), (3.1)–(3.3) is the Skorokhod embedding of the standard (scaled) random walk into the standard Brownian motion. In [4] the author proved that in this case

$$\mathbb{E}\left(\sup_{0 \leq t \leq T} |X_{[nt/T]}^{(n)} - X_t|\right) = O(n^{-1/4}(\ln n)^{3/4}).$$

However, it is not clear whether this estimate is sharp. In general, the Skorokhod embedding technique does not allow to get error estimates of order $O(n^{-1/4})$. For details see Remark 3.8 in [4].

4. APPLICATIONS TO OPTIMAL STOPPING

In this section we will discuss the applications of Theorem 2.2 to finite horizon optimal stopping problems where the underlying process is the diffusion X .

Let $f : [0, T] \times \mathbb{R} \rightarrow \mathbb{R}$ be a Lipschitz continuous function. Denote by \mathcal{T}_T the set of all stopping times with respect to the Brownian filtration \mathcal{F}_t , $t \geq 0$ with values in $[0, T]$. Consider the optimal stopping problem

$$(4.1) \quad V = \sup_{\tau \in \mathcal{T}_T} \mathbb{E}(f(\tau, X_\tau)).$$

We use Theorem 2.2 in order to calculate numerically V and the optimal control τ^* . Let us remark that the optimal control exists since we deal with a continuous payoff in the Brownian framework (for details see [6] Chapter I).

We start with introducing the appropriate optimal stopping problem for the correlated random walks. For any n denote by \mathcal{T}_n the (finite) set of all stopping times with respect to filtration $\sigma\{\xi_1^{(n)}, \dots, \xi_k^{(n)}\}$, $k = 0, \dots, n$ which take on values in the set $\{0, 1, \dots, n\}$. Set,

$$(4.2) \quad V_n = \max_{\tau \in \mathcal{T}_n} \mathbb{E}_n \left(f(\tau h, X_\tau^{(n)}) \right).$$

By using standard dynamical programming for optimal stopping (see [6] Chapter I) we can calculate V_n and the optimal stopping time τ_n^* by the following dynamical programming. Define the functions (recall (2.5))

$$J_k^{(n)} : \{x + \sqrt{h}\{-k, 1-k, \dots, 0, 1, \dots, k\}\} \times \{-1, 1\} \rightarrow \mathbb{R}, \quad k = 0, 1, \dots, n$$

by the following recursion.

$$J_n^{(n)}(z, y) = f(T, z)$$

$$\text{and for } k = 0, 1, \dots, n-1, \quad J_k^{(n)}(z, y) = \max \left(f(kh, z), \right.$$

$$\left. \sum_{i=1}^2 \frac{1}{2} \left(1 + (-1)^i \frac{\alpha' y + \sqrt{h} \mu(z)}{1 + \alpha} \right) f((k+1)h, z + (-1)^i \sqrt{h}) \right)$$

where $\alpha' = \frac{\sigma^2(z - (-1)^y \sqrt{h}) - 1}{2}$ and $\alpha = \frac{\sigma^2(z) - 1}{2}$. We get that

$$V_n = J_0^{(n)}(x, 1)$$

and the stopping time given by

$$(4.3) \quad \tau_n^* = n \wedge \min \left\{ k : J_k^{(n)}(X_k^{(n)}, \xi_k^{(n)}) = f(kh, X_k^{(n)}) \right\}$$

is the optimal stopping time. Namely,

$$(4.4) \quad V_n = \mathbb{E}_n \left(f(\tau_n^* h, X_{\tau_n^*}^{(n)}) \right).$$

Next, recall the Skorokhod embedding given by (3.1)–(3.3). These formulas allow to redefine the random times τ_n^* , $n \in \mathbb{N}$ on the Brownian probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a given n consider the random time $T \wedge \theta_{\tau_n^*}^{(n)}$. Clearly, this random time is a stopping time with respect to the Brownian filtration, i.e. $T \wedge \theta_{\tau_n^*}^{(n)} \in \mathcal{T}_T$. From practical point of view, by calculating the functions $J_k^{(n)}$, $k \leq n$ defined above, and observing the diffusion X we can calculate explicitly the stopping time $T \wedge \theta_{\tau_n^*}^{(n)}$.

The following result says that V_n is a good approximation of V . Moreover, the stopping time $T \wedge \theta_{\tau_n^*}^{(n)}$ is asymptotically optimal for the primal optimal stopping problem given by (4.1).

Theorem 4.1. *For any $\epsilon > 0$*

$$(4.5) \quad \lim_{n \rightarrow \infty} n^{1/4-\epsilon} |V_n - V| = 0$$

and

$$(4.6) \quad \lim_{n \rightarrow \infty} n^{1/4-\epsilon} \left(V - \mathbb{E} \left(f(T \wedge \theta_{\tau_n^*}^{(n)}, X_{T \wedge \theta_{\tau_n^*}^{(n)}}) \right) \right) = 0.$$

Proof. First we make some preparations. Again, in view of the Skorokhod embedding given by (3.1)–(3.3) we consider the random variables $\xi_k^{(n)}$, $k \leq n$ on the Brownian probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For a given n denote by \mathcal{S}_n the set of all stopping times with respect to the filtration $\{\mathcal{F}_{\theta_k^{(n)}}\}_{k=0}^n$ with values in the set $\{0, 1, \dots, n\}$. Clearly, $\mathcal{T}_n \subset \mathcal{S}_n$, where \mathcal{T}_n is defined before (4.2).

From (3.5)–(3.6) it follows that

$$\mathbb{P}(X_{k+1}^{(n)}, \xi_{k+1}^{(n)} | \mathcal{F}_{\theta_k^{(n)}}) = \mathbb{P}(X_{k+1}^{(n)}, \xi_{k+1}^{(n)} | \xi_1^{(n)}, \dots, \xi_k^{(n)}), \quad k = 0, 1, \dots, n-1.$$

Hence, by applying the dynamical programming for optimal stopping we have

$$(4.7) \quad V_n = \max_{\tau \in \mathcal{T}_n} \mathbb{E} \left(f(\tau h, X_\tau^{(n)}) \right) = \sup_{\tau \in \mathcal{S}_n} \mathbb{E} \left(f(\tau h, X_\tau^{(n)}) \right), \quad n \in \mathbb{N}.$$

Next, we derive two estimates that will be used in the proof of the theorem. Using (3.18) we obtain that for any $m > 1$

$$(4.8) \quad \mathbb{E} \left(\max_{0 \leq k \leq n-1} \max_{\theta_k^{(n)} \leq t \leq \theta_{k+1}^{(n)}} |X_t - X_{\theta_k^{(n)}}|^m \right) \leq O(1) \sum_{k=0}^{n-1} \mathbb{E} \left((\theta_{k+1}^{(n)} - \theta_k^{(n)})^{m/2} + (\theta_{k+1}^{(n)} - \theta_k^{(n)})^m \right) = nO(h^{m/2}) = O(h^{m/2-1})$$

where the last equality follows from (3.14). Similarly, from (3.20)

$$(4.9) \quad \mathbb{E} \left(\max_{\theta_n^{(n)} \wedge T \leq t \leq \theta_n^{(n)} \vee T} |X_t - X_{\theta_n^{(n)} \wedge T}|^m \right) \leq O(1) \mathbb{E} \left(|\theta_n^{(n)} - T|^{m/2} + |\theta_n^{(n)} - T|^m \right) = O(h^{m/4-1}), \quad \forall m > 2.$$

Next we complete the proof in two steps.

First Step: In this step we show that for any $\epsilon > 0$

$$(4.10) \quad V < V_n + O(n^{1/4-\epsilon}).$$

Let $\tau^* \in \mathcal{T}$ satisfies

$$(4.11) \quad V = \mathbb{E}(f(\tau^*, X_{\tau^*})).$$

Define the random time $\zeta_n = n \wedge \min\{k : \theta_k^{(n)} \geq \tau^*\}$. It is straightforward to verify that $\zeta \in \mathcal{S}_n$. Thus from (4.7) and (4.11) it follows

$$V - V_n \leq \mathbb{E} \left(f(\tau^*, X_{\tau^*}) - f(\zeta_n h, X_{\zeta_n}^{(n)}) \right) \leq$$

from the definition of ζ_n and the fact that f is Lipschitz

$$\begin{aligned} &\leq O(1)\mathbb{E}\left(\max_{0\leq k\leq n-1}\sup_{\theta_k^{(n)}\leq t\leq\theta_{k+1}^{(n)}}|t-(k+1)h|+|T-\theta_n^{(n)}|\right)+ \\ &\quad O(1)\mathbb{E}\left(\max_{0\leq k\leq n-1}\sup_{\theta_k^{(n)}\leq t\leq\theta_{k+1}^{(n)}}|X_t-X_{k+1}^{(n)}|\right)+ \\ &\quad O(1)\mathbb{E}\left(\max_{\theta_n^{(n)}\wedge T\leq t\leq\theta_n^{(n)}\vee T}|X_t-X_n^{(n)}|\right)= \end{aligned}$$

by simple rearranging we obtain

$$\begin{aligned} &= O(1)\mathbb{E}\left(h+\max_{0\leq k\leq n}|\theta_k^{(n)}-kh|\right)+ \\ &\quad O(1)\mathbb{E}\left(\max_{0\leq k\leq n}|X_{\theta_k^{(n)}}-X_k^{(n)}|\right) \\ &+ O(1)\mathbb{E}\left(\max_{0\leq k\leq n-1}\sup_{\theta_k^{(n)}\leq t\leq\theta_{k+1}^{(n)}}|X_t-X_{\theta_k^{(n)}}|\right)+ \\ &\quad O(1)\mathbb{E}\left(\max_{\theta_n^{(n)}\wedge T\leq t\leq\theta_n^{(n)}\vee T}|X_t-X_{\theta_n^{(n)}\wedge T}|\right)= \end{aligned}$$

by applying (3.17), (3.20), (4.8)–(4.9) and the Jensen inequality

$$\begin{aligned} &= O(h)+\left(\sum_{k=0}^{n-1}\mathbb{E}(|\theta_k^{(n)}-kh|^m)\right)^{1/m}+ \\ &\quad +O(\sqrt{h})+O(h^{1/2-1/m})+O(h^{1/4-1/m})=O(h^{1/4-1/m}), \quad \forall m>4. \end{aligned}$$

This completes the proof of (4.10).

Second Step: Recall the stopping time τ_n^* given by (4.3). Since $T\wedge\theta_{\tau_n^*}^{(n)}\in\mathcal{T}$ then

$$V\geq\mathbb{E}\left(f(T\wedge\theta_{\tau_n^*}^{(n)},X_{T\wedge\theta_{\tau_n^*}^{(n)}})\right).$$

Thus in view of (4.10) in order to prove (4.5)–(4.6) it remains to show that for any $\epsilon>0$

$$(4.12) \quad \left|V_n-\mathbb{E}\left(f(T\wedge\theta_{\tau_n^*}^{(n)},X_{T\wedge\theta_{\tau_n^*}^{(n)}})\right)\right|=O(n^{1/4-\epsilon}).$$

From (4.4) and the fact that f is Lipschitz it follows

$$\begin{aligned} &\left|V_n-\mathbb{E}\left(f(T\wedge\theta_{\tau_n^*}^{(n)},X_{T\wedge\theta_{\tau_n^*}^{(n)}})\right)\right|\leq\mathbb{E}\left(\left|f(\tau_n^*h,X_{\tau_n^*}^{(n)})-f(T\wedge\theta_{\tau_n^*}^{(n)},X_{T\wedge\theta_{\tau_n^*}^{(n)}})\right|\right) \\ &= O(1)\mathbb{E}\left(\max_{0\leq k\leq n}|\theta_k^{(n)}-kh|\right)+O(1)\mathbb{E}\left(\max_{0\leq k\leq n}|X_{kh}-X_{\theta_k^{(n)}}|\right)+ \\ &\quad O(1)\mathbb{E}\left(\max_{\theta_n^{(n)}\wedge T\leq t\leq\theta_n^{(n)}\vee T}|X_t-X_{\theta_n^{(n)}\wedge T}|\right)=O(h^{1/4-1/m}), \quad \forall m>4. \end{aligned}$$

The last equality follows by using the same arguments as in the first step. This completes the proof of Theorem 4.1. \square

REFERENCES

- [1] D.S Clark, *Short Proof of a Discrete Gronwall Inequality*, Discrete Applied Mathematics. **16** (1987), 279–281.
- [2] S.Ethier and T Kurtz, *Markov Processes. Characterization and Convergence*, Wiley, New York, 1986.
- [3] U.Gruber and M.Schweizer, *A Diffusion Limit for Generalized Correlated Random Walks*, J.Appl. Probab. **43** (2006), 60–73.
- [4] Yu.Kifer, *Error estimates for binomial approximations of game options*, Ann. Appl. Probab. **16** (2006), 984–1033.

- [5] P.E.Kloeden and E.Platen, *Numerical Solution of Stochastic Differential Equations*, volume 23. Springer, 1992.
- [6] G.Peskir and A.N.Shiryaev, *Optimal Stopping and Free-Boundary Problems*, Lectures in Mathematics, ETH Zurich, Birkhauser, (2006).

DEPARTMENT OF STATISTICS, HEBREW UNIVERSITY OF JERUSALEM, ISRAEL.
E.MAIL: YAN.DOLINSKY@MAIL.HUJI.AC.IL